UNITED STATES DEPARTMENT OF COMMERCE
Patent and Trademark Office
Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.
08/970,889	11/14/97	BERGEN J	SAR12228

LM02/0106

WILLIAM J BURKE
LAW AND PATENT OPERATION
SARNOFF CORPORATION
201 WASHINGTON ROAD CN 5300
PRINCETON NJ 08543-5300

EXAMINER	
PADMANABHAN, M	
ART UNIT	**PAPER NUMBER**
2772	8

DATE MAILED: 01/06/00

Please find below and/or attached an Office communication concerning this application or proceeding.

Commissioner of Patents and Trademarks

Office Action Summary

Application No.	Applicant(s)
08/970,889	Russell et al.
Examiner	**Group Art Unit**
Mano Padmanabhan	2772

☒ Responsive to communication(s) filed on *Oct 12, 1999* .

☒ This action is **FINAL**.

☐ Since this application is in condition for allowance except for formal matters, **prosecution as to the merits is closed** in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11; 453 O.G. 213.

A shortened statutory period for response to this action is set to expire *3* month(s), or thirty days, whichever is longer, from the mailing date of this communication. Failure to respond within the period for response will cause the application to become abandoned. (35 U.S.C. § 133). Extensions of time may be obtained under the provisions of 37 CFR 1.136(a).

Disposition of Claims

☒ Claim(s) *1-11, 13, 14, and 17-26* is/are pending in the application.

Of the above, claim(s) _____ is/are withdrawn from consideration.

☐ Claim(s) _____ is/are allowed.

☒ Claim(s) *1-11, 13, 14, and 17-26* is/are rejected.

☐ Claim(s) _____ is/are objected to.

☐ Claims _____ are subject to restriction or election requirement.

Application Papers

☐ See the attached Notice of Draftsperson's Patent Drawing Review, PTO-948.

☐ The drawing(s) filed on _____ is/are objected to by the Examiner.

☐ The proposed drawing correction, filed on _____ is ☐approved ☐disapproved.

☐ The specification is objected to by the Examiner.

☐ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. § 119

☐ Acknowledgement is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d).

☐ All ☐ Some* ☐ None of the CERTIFIED copies of the priority documents have been

☐ received.

☐ received in Application No. (Series Code/Serial Number) _____ .

☐ received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

*Certified copies not received: _____ .

☐ Acknowledgement is made of a claim for domestic priority under 35 U.S.C. § 119(e).

Attachment(s)

☒ Notice of References Cited, PTO-892

☐ Information Disclosure Statement(s), PTO-1449, Paper No(s). _____

☐ Interview Summary, PTO-413

☐ Notice of Draftsperson's Patent Drawing Review, PTO-948

☐ Notice of Informal Patent Application, PTO-152

--- SEE OFFICE ACTION ON THE FOLLOWING PAGES ---

DETAILED ACTION

Status of claims

Claims 1-11, 13-14, 17-26 are in the Application.

Claims 12, 15, and 16 are canceled.

Claims 1-11, 13-14, 17-26 are rejected.

Applicants arguments and amendments filed on September 18, 1999 in response to the office action mailed on July 12, 1999 have been fully considered.

Claim Rejections - 35 USC § 103

Applicants amendments to the claims to overcome the 35 USC 103 rejections made in the previous office action has necessitated the following new rejections. The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

1. Claims 1-3, 11, and 21-23 are rejected under 35 U.S.C. 103(a) as being unpatentable over Adelson (U.S. Patent 5,706,417) in view of Yeo et al. (U.S. Patent 5,821,945), and Shibata et al. ("Content-Based structuring of video information": 0-8186-7436-9/96, 1996 IEEE).

Claim 1 lays claim to a method of representing video information comprising the steps of segmenting a video stream into scenes, each scene into frames including a key frame, and also dividing scenes into at least one background and at least one foreground layer using intra-scene

motion analysis, and storing content-related appearance attributes or mosaic representations in a database.

Claim 2 adds to claim 1 the step of computing, and storing content-related appearance attributes for the background and foreground layers.

Adelson teaches that a layer exists for each object, set of objects, or portion of an object in the image having a motion vector significantly different from any other object in the image (Col.2: lines 45-47). He also teaches combining the foreground and background images to produce a video image (Col.2: lines 15-21; Col.6: lines 50-55). Adelson also teaches content related appearance attributes for each layer with the use of intensity map, attenuation map, velocity map, and delta map (Col. 2: lines 50-67), and implicitly teaches storing these attributes in a database. Adelson does not teach segmenting a video stream into scenes, and scenes into frames including a key frame, and the use of intra-scene motion analysis. Yeo discloses dividing the sequence into equal length segments, denoting the first frame of each segment as its key frame (Col.1: lines 34-38), and also teaches classifying a long video sequence into story units (Col.1: lines 47-50). Shibata teaches segmenting a video sequence, with individual video frames being the smallest unit of any segment. He also teaches the use of a basic segment which is a collection of video frames having the same vector expressions, assuming a collection of basic segments as the initial layer, and creating new layers by adding a segment to the previously processed layer, thus teaching a method for providing background mosaic, and intra-scene

motion analysis. Hence it would be obvious to one skilled in the art at the time the invention was made to segment a video stream into scenes containing video frames with a key frame for each scene, as this will provide an effective means of browsing the video content.

Claim 3 adds to claim 2 the steps of storing the scenes in a mass storage unit, and retrieving scenes associated with an attribute.

Adelson teaches the use of video tape player, laser disc player as data source for image pixel data (Col.4: lines 2-7, 16-20). This implicitly teaches using mass storage unit to store data representing the scenes. Adelson also teaches having various maps for the various attributes (Col.2: lines 55-67; Col.5: lines 9-17), and retrieving data easily to reconstruct an image, based on the required image (Col.6: lines 30-47).

Claim 11 adds to claim 1 the steps of storing ancillary information related to layers or frames.

Adelson teaches the use of optional maps, including a contrast change map and a blur map for each layer (Col.3: lines 6-14).

Claim 21 is a claim for a computer readable medium that implements the method as claimed in claim 1 and hence is rejected for the same reasons.

Claim 22 is a claim for a computer readable medium that implements the method as claimed in claim 2 and hence is rejected for the same reasons.

Claim 23 is a claim for a computer readable medium that implements the method as claimed in claim 3 and hence is rejected for the same reasons.

2. Claims 4 and 24 are rejected under 35 U.S.C. 103(a) as being unpatentable over Adelson (U.S. Patent 5,706,417) in view of Yeo et al. (U.S. Patent 5,821,945) and Shibata et al, as applied to claims 1 and 22 respectively, and further in view of Jaillon et al. ("Image Mosaicing Applied to Three-Dimensional Surfaces": Jaillon et al.; 1051-4651/94 - 1994 IEEE).

Claim 4 adds to claim 1 the limitation that the mosaic representation is one of a two dimensional, a three dimensional, and a network of mosaics.

Jaillon teaches aligning and combining images or other mosaics to form a mosaic. Hence it would be obvious to one skilled in the art at the time the invention was made to combine various layers/images to generate a mosaic representation as this will provide the user greater flexibility in altering the image scene to suit their needs.

Claim 24 is a claim for a computer readable medium that implements the method as claimed in claim 4 and hence is rejected for the same reasons.

3. Claims 5-8 are rejected under 35 U.S.C. 103(a) as being unpatentable over Adelson (U.S. Patent 5,706,417) in view of Yeo et al. (U.S. Patent 5,821,945), and Shibata et al, as applied to claim 2, and further in view of Jaillon et al. ("Image Mosaicing Applied to Three-Dimensional Surfaces": Jaillon et al.; 1051-4651/94 - 1994 IEEE).

Claim 5 adds to claim 2 the steps of generating an image pyramid for a layer, filtering such that each subband is associated with feature maps, and integrating feature maps to produce attribute pyramid subbands, which comprise content-based appearance attribute subband associated with a corresponding image pyramid subband.

Adelson discloses the use of subbands to encode images (Col.1: lines 20-24). Adelson also teaches the feature maps associated with each layer (Col.2: lines 55-67; Col.5: lines 9-17), and integrating the feature maps to reconstruct an image (Col.6: lines 30-47). Adelson and Yeo fail to teach image pyramids. Jaillon teaches the use of image pyramid framework in the alignment process, and converting the input image and the mosaic into Laplacian image pyramids, and applying the alignment to all levels within the respective pyramids. Hence it would be obvious to one skilled in the art at the time the invention was made to use the image pyramid in each layer in order to achieve better alignment and reproduction of the image.

Claim 6 adds to claim 5 the limitation that the attribute comprises at least one of luminance, chrominance, and texture.

Adelson discloses the use of intensity map, depth map, blur map, contract change map (Col.2: lines 55-67; Col.5: lines 9-17).

Claim 7 adds to claim 5 the step of rectifying the feature maps associated with each subband.

Adelson discloses the use of delta map, which is essentially an additive error map, which provides correction data for any changes in the image over time which can not be accounted for by the other maps.

Claim 8 adds to claim 5 the step of collapsing the attribute pyramid subbands to produce a content-based appearance attribute.

Yeo teaches that the lower levels of the hierarchy can be based on visual cues, while the upper levels allow criteria that reflect semantic information associated (Col.5: lines 48-52), the nodes capturing the contents of a video, while the edges capture its structure. Yeo also teaches a tree hierarchy that permits the user to have a coarse-to-fine view of the entire video sequences based on the level of the nodes (Col.4: lines 30-35), the nodes capturing the core contents of the video while the edges capture its structure (Col.5: lines 40-43). Hence it would be obvious to one skilled in the art at the time the invention was made to collapse the attribute pyramid subbands to produce a content-based appearance attribute since this will offer a browsing structure that closely resembles human perception and understanding.

4. Claims 9-10 are rejected under 35 U.S.C. 103(a) as being unpatentable over Adelson (U.S. Patent 5,706,417) in view of Yeo et al. (U.S. Patent 5,821,945), and Shibata et al., as applied to claim 2, and further in view of Barber et al. (U.S. Patent 5,751,286).

Claim 9 adds to claim 2 the steps of receiving a request matching a desired content-related appearance attribute, and retrieving at least one layer matching the request.

Adelson teaches a method of retrieving data representing layers, each layer comprising a series of maps, to reconstruct an image. Barber teaches a method of building a visual query by image content, and retrieving database images with features that correspond to the selected image characteristics (Col.2: line 64 - Col.3: line 8). Hence it would be obvious to one skilled in the art at the time the invention was made to query the database by content-related appearance attribute, and retrieve layers that match the attribute, in order to reconstruct the image as desired, as such an approach will save database storage requirements.

Claim 10 adds to claim 9 the steps of identifying a query type as being one of luminance, chrominance, and texture type, and a query specification as being a desired property of the query type, and selecting a filter type and calculating the appearance attribute based on filter type and desired property.

Barber discloses a query construction interface with a hierarchical selection windows for each of image color, shapes, textures, category, which may include keywords, text or conditions (Col.3: lines 22-34). Barber also teaches filtering the masks in the current image by the category

code, establishing the set of masks that will be analyzed with respect to the image characteristic values (Col.12: lines 1-5). Barber also teaches computing positional feature score that compares the area's similarity to the image areas (Col.14: lines 40-60). Hence it would be obvious to one skilled in the art at the time the invention was made to use a query type to chose the parameter, and specification to specify a desired property for the parameter, as this would facilitate retrieving only the layers that match the selection criteria, and hence would increase the speed of rendering the image.

5. Claims 13-14 are rejected under 35 U.S.C. 103(a) as being unpatentable over Adelson (U.S. Patent 5,706,417) in view of Yeo et al. (U.S. Patent 5,821,945) and Shibata et al., as applied to claim 1, and further in view of Zhang et al. (U.S. Patent 5,635,982).

Claim 13 adds to claim 1 the steps of generating a descriptor vector, and generating a scene cut indicium in response to calculated differences between descriptive vectors of successive frames exceeding a threshold.

Adelson teaches generating an intensity map, an attenuation map, a velocity map, and a delta map for each layer. Zhang teaches calculating the differences between consecutive video frames based on the selected difference metric, and defining a cut if the values exceed a threshold value (Col.7: lines 1-10; Col.8: lines 5-15). Hence it would be obvious to one skilled in the art at the time the invention was made to generate a scene cut if the calculated differences between

descriptive vectors exceeded a threshold value, as this would minimize the calculation needed to detect scene cuts.

Claim 14 adds to claim 1 the steps of generating a descriptor vector and a threshold for it in the first pass, and calculating the difference between the frames and generating a scene cut indicium in the second pass, if the difference exceeds the threshold value.

Zhang teaches a multi-pass approach, wherein the prospective segment boundaries are determined in the first pass, by comparing against a threshold value. This implies the use of a descriptor vector to define a frame, such that they can be compared against a threshold value. Zhang teaches using the second pass to locate all boundaries (scene cuts). Zhang also teaches using the multi-pass approach to apply different difference metrics in different passes (Col.6: lines 20-64), and teaches defining cuts based on the differences in the difference metrics (Col.8: lines 5-15). Hence it would be obvious to one skilled in the art at the time the invention was made to use two passes as described in this claim to compute the attribute value, as this would provide more accurate values for the attribute.

8. Claims 17-20 are rejected under 35 U.S.C. 103(a) as being unpatentable over Barber et al. (U.S. Patent 5,751,286) in view of Yeo et al. (U.S. Patent 5,821,945) and Shibata et al.

Claim 17 claims a method for browsing a video program comprising a plurality of scenes that contain frame(s), comprising the steps of providing a database comprising attribute information, formulating a query utilizing the attribute information, and searching and retrieving video frames that substantially match the query criterion.

Barber teaches a query facility which builds a visual query by image content, and also teaches a query engine that interprets the query, and returns database images with features that correspond to the selected criteria (Col.2: line 64 - Col.3: line 8). Barber does not teach the notion of a representative video frame for a video scene. Yeo discloses a method for content-based video browsing, containing a video database, and sets of key frames that have associated attributes, the key frames representing the long sequence of related shots (Col.2: lines 35-45). Yeo also teaches the use of Rframes (representative frames) to organize the visual contents of the video clips (Col.1: lines 30-65). Shibata teaches segmenting a video sequence, with individual video frames being the smallest unit of any segment. He also teaches the use of a basic segment which is a collection of video frames having the same vector expressions, assuming a collection of basic segments as the initial layer, and creating new layers by adding a segment to the previously processed layer, thus teaching a method for providing background mosaic, and intra-scene motion analysis. Hence it would be obvious to one skilled in the art at the time the invention was made to build a query to retrieve the representative frames, as this would be a faster way to identify areas of interest before retrieving all the related frames.

Claim 18 adds to claim 17 the steps of selecting a query type, query specification, and computing a multi-dimensional feature vector.

Barber teaches query specification for image characteristics (query type) (Col.13: lines 44-53). Barber also teaches calculating a positional feature score combining features and positional similarity for each of the areas selected in the query (Col.15: lines 40-61).

Claim 19 adds to claim 18 the limitation of selecting a query specification by identifying a portion of the displayed image, and the feature vector is calculated based on query type and the identified image portion.

Barber teaches specification in a query of image characteristics that occur in some area or areas of the image (Col.13: lines 45-52).

Claim 20 adds to claim 19 the steps of formatting and transmitting the identified video frames.

Barber teaches returning the images with the best scores in response to a query (Col.14: lines 65-67).

3. Claims 25 - 26 are rejected under 35 U.S.C. 103(a) as being unpatentable over Adelson (U.S. Patent 5,706,417) in view of Yeo et al. (U.S. Patent 5,821,945), and Shibata et al, as

applied to claim 22, and further in view of Jaillon et al. ("Image Mosaicing Applied to Three-Dimensional Surfaces": Jaillon et al.; 1051-4651/94 - 1994 IEEE).

Claim 25 is a claim for a computer readable medium that implements the method as claimed in claim 5 and hence is rejected for the same reasons.

Claim 26 is a claim for a computer readable medium that implements the method as claimed in claim 6 and hence is rejected for the same reasons.

Response to Arguments: Applicant's arguments with respect to claims 1-11, 13-14, 17-26 have been considered but are moot in view of the new ground(s) of rejection.

Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event,

however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Conclusion

Any response to this correspondence should be mailed to:

Box AF

Commissioner of Patents and Trademarks

Washington, D.C. 20231

or faxed to:

(703) 305-9051, (for formal communications; please mark "EXPEDITED PROCEDURE")

Or:

(703) 308-6606 (for informal or draft communications, please label "PROPOSED" or "DRAFT")

Hand-delivered responses should be brought to Crystal Park II, 2021 Crystal Drive, Arlington. VA., Sixth Floor (Receptionist).

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Mano Padmanabhan whose telephone number is (703) 306-2903. She can normally be reached Monday-Thursday from 6:30am-5:00pm.

If attempts to reach the examiner are unsuccessful, the examiner's supervisor, Mark Powell, can be reached on (703) 305-9703.

Any inquiry of a general nature or relating to the status of this application should be directed to the Group receptionist whose telephone number is (703) 305-3900.

Mano Padmanabhan

December 28, 1999

MARK R. POWELL
SUPERVISORY PATENT EXAMINER
GROUP 2700